STORET Import Module (SIM)

Web Version

User Guide and Reference Manual

U.S. Environmental Protection Agency
November 2004

Table of Contents

Introducing SIM

The STORET Import Module (SIM) is a Web program that helps users load data into STORET, an EPA-maintained database of ambient environmental data. STORET is a distributed system that allows individual agencies to manage their own data at a local level and to share their data with others via a national data warehouse.

To use SIM, you must perform the following tasks:

- Organize your data into delimited text files
- Establish SIM import configurations that describe the format of your text files
- Import your text files into SIM and migrate the data that pass validation into STORET.

This user guide and reference manual will

- Guide you in creating your delimited text files
- Teach you how to create a SIM import configuration
- Walk you through the process of using SIM to load data into STORET.

Note: To use SIM successfully, you need a clear understanding of STORET. Before you begin to use SIM, you should use the STORET Web Registration Application to establish all relevant metadata defaults in STORET.

Using This Guide

This guide is designed to provide you with "hands on" instruction for using SIM. Each lesson focuses on different functions included in SIM, describes the functions and associated concepts, and provides step-by-step instructions for using the tools. The step-by-step instructions are called out in boxes, which are separate from the text (see box below).

1. This is an example of a box that contains instructions for using the guide with the sample data.

For assistance, contact the Central Data Exchange (CDX) Help Desk at

- EPACDX@CSC.COM
- 888-890-1995

Formatting Your Data

Before using SIM, you must organize your data into delimited text files (ASCII flat files). Most organizations use common software products, such as Excel, Access, or Lotus 1-2-3, to create data tables that can be saved and exported as delimited text files. Larger organizations may produce these files as exports from existing data management or laboratory information management systems (LIMS). SIM accepts two different file types:

1. Station Descriptions—Locations where field measurements are made or samples are collected

2. Field Measurements and Samples—Results of field measurements or sample analyses.

The data in the two file types complement each other and the order in which you migrate data files to STORET is important. Station descriptions should always be loaded before field measurements and samples. This ensures that the key data elements will be registered in STORET and are ready for results to be associated with them.

For ASCII flat files, the data elements or fields in each file type are organized in a tabular format where the column delimiter is a tab ( ), pipe (|), tilde (~), or comma (,).[1] Additionally, a subdelimiter (\) is available for columns that can accept multiple values.[2] The required and optional elements for each file type are described in Appendix A. Many of these data elements have a list of valid values that you must select from. The order in which you place the data elements in your delimited text files is not important because SIM allows you to define the order before you import your file.

Note: Do not include a header row in the delimited text files that you create, and do not use text qualifiers for text values (e.g., “”), because header rows and text qualifiers can result in errors in your import files.

[1] If any of these characters appear in your data, you should not use that character as a delimiter because there is no way for SIM to tell which is the delimiter and which is simply part of the data.

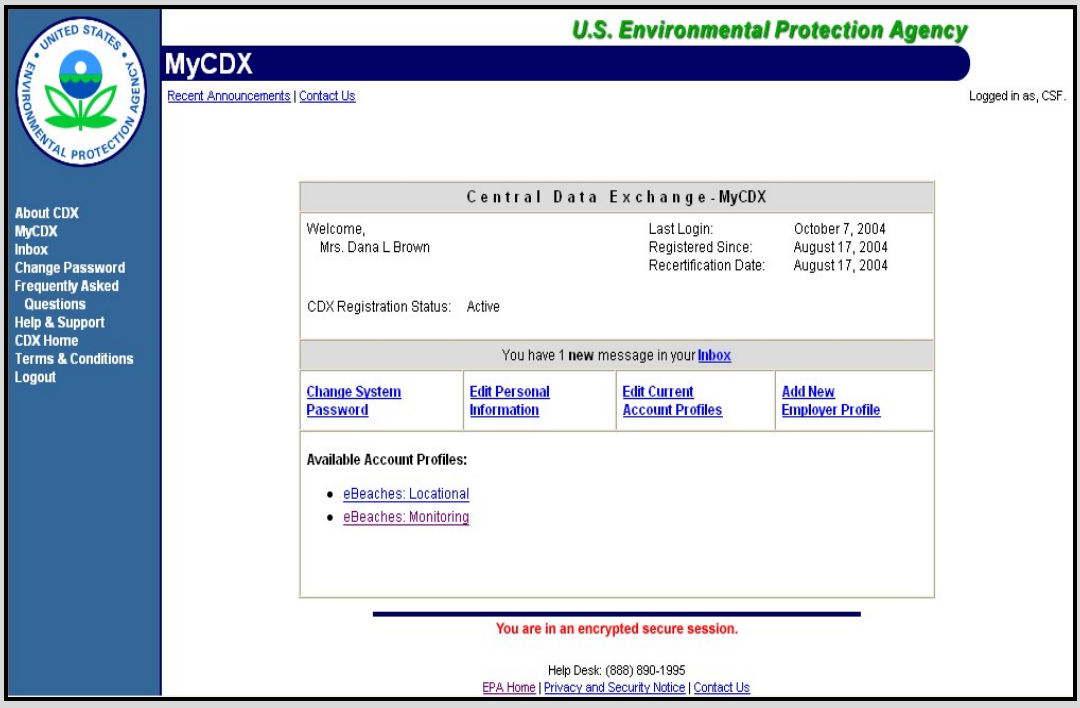
[2] Certain fields, such as Project ID in the “Field Measurements and Non-Biological Samples” section, can accept multiple values in a single column and row of the import file. See Appendix A for a list of columns that support loading multiple values.

Lesson I. Defining an Import Configuration

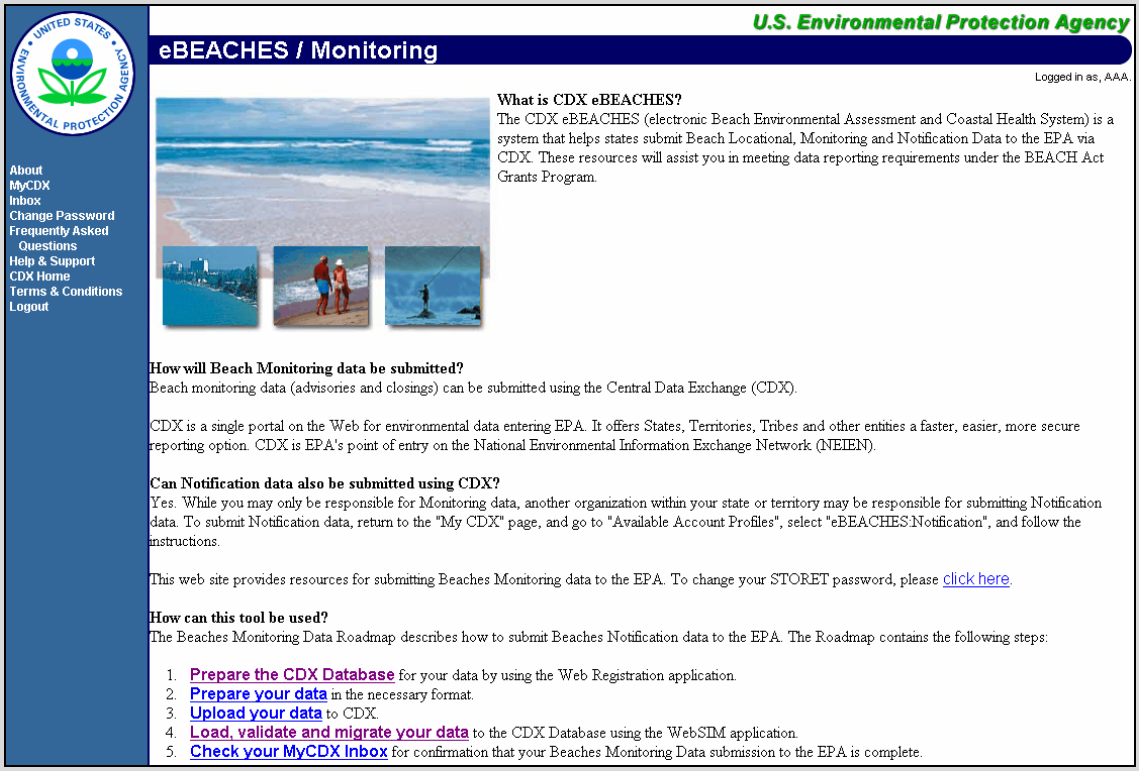
The following lesson will show you how to create a custom import configuration that will describe the format of your text files.

1. First, log in to CDX and access "MyCDX" from http://cdx.epa.gov/warning.asp. Read the warning notice and privacy statement and click the link at the bottom of the page to advance.

Next, click the link for users who already have a CDX ID and password (the second choice in the bulleted list). Enter your CDX ID and password and click **LOGIN**. This will give you access to the MyCDX page. You should have access to the monitoring role to transmit beach monitoring data. If monitoring is not an option under the MyCDX Profiles, call the CDX Help Desk. Click on **eBEACHES: Monitoring** near the bottom of the page.



2. Click **Load, validate, and migrate your data** (item 4 near the bottom of the page).

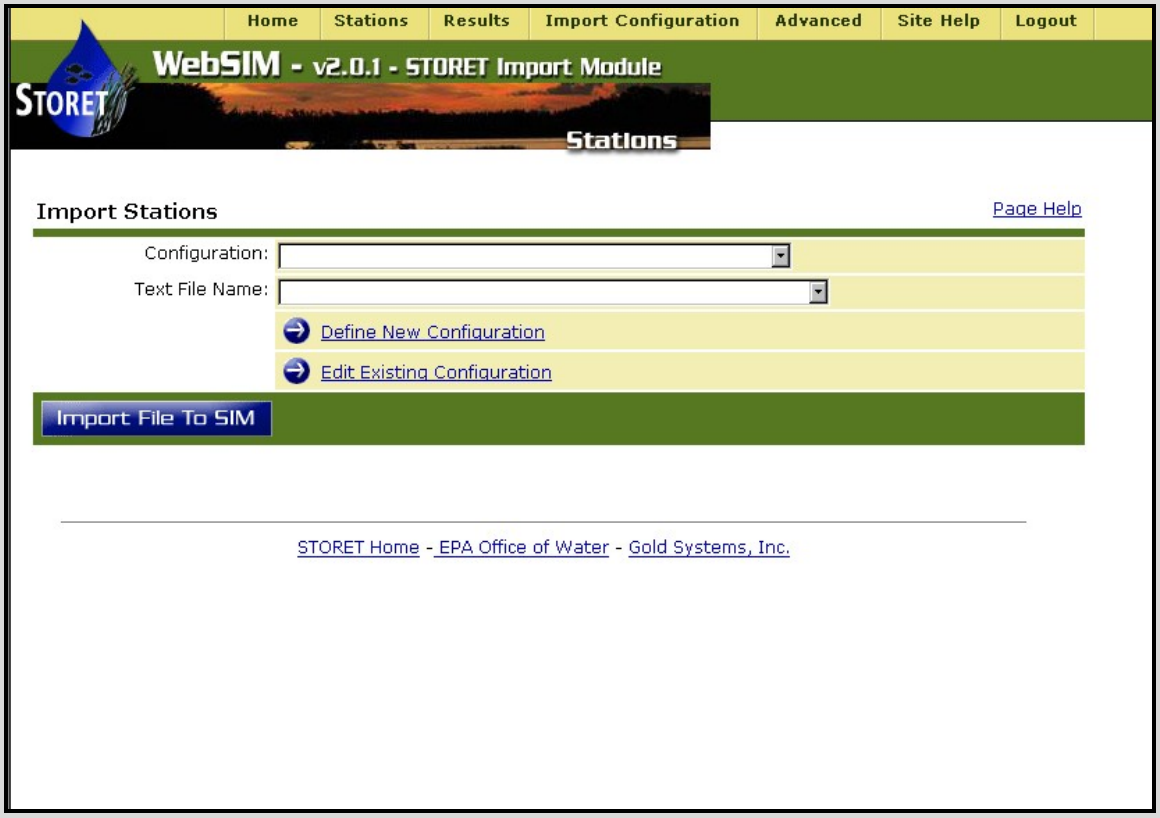


3. Click the **WebSIM application** link. Next, log on to WebSIM by clicking **Log on to SIM**. Enter the appropriate user name and password.

4. You can either start a new import of data or access an existing import that you may have brought into SIM but have yet to migrate to the STORET database. For this lesson, click on **Stations** on the toolbar, then click on **New Import.**

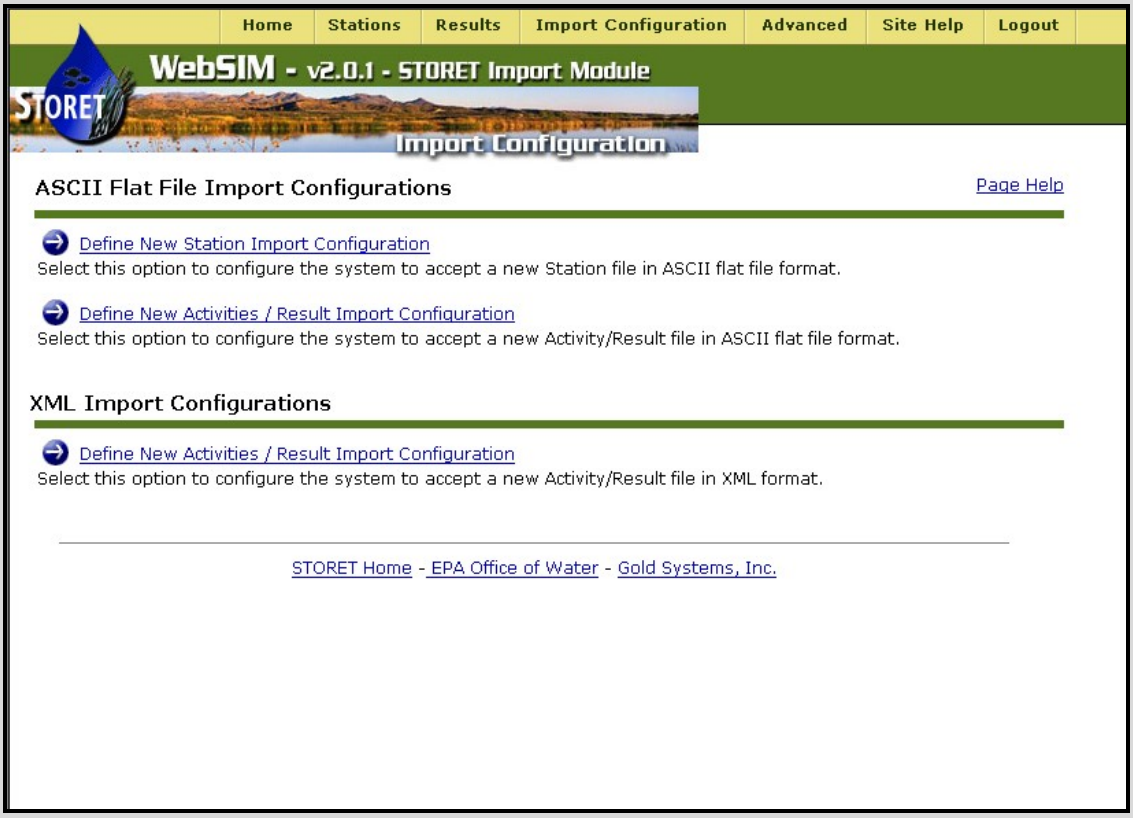
5. This menu allows you to specify the import configuration and the path to the file that you want to import.

   Click **Define New Configuration.**

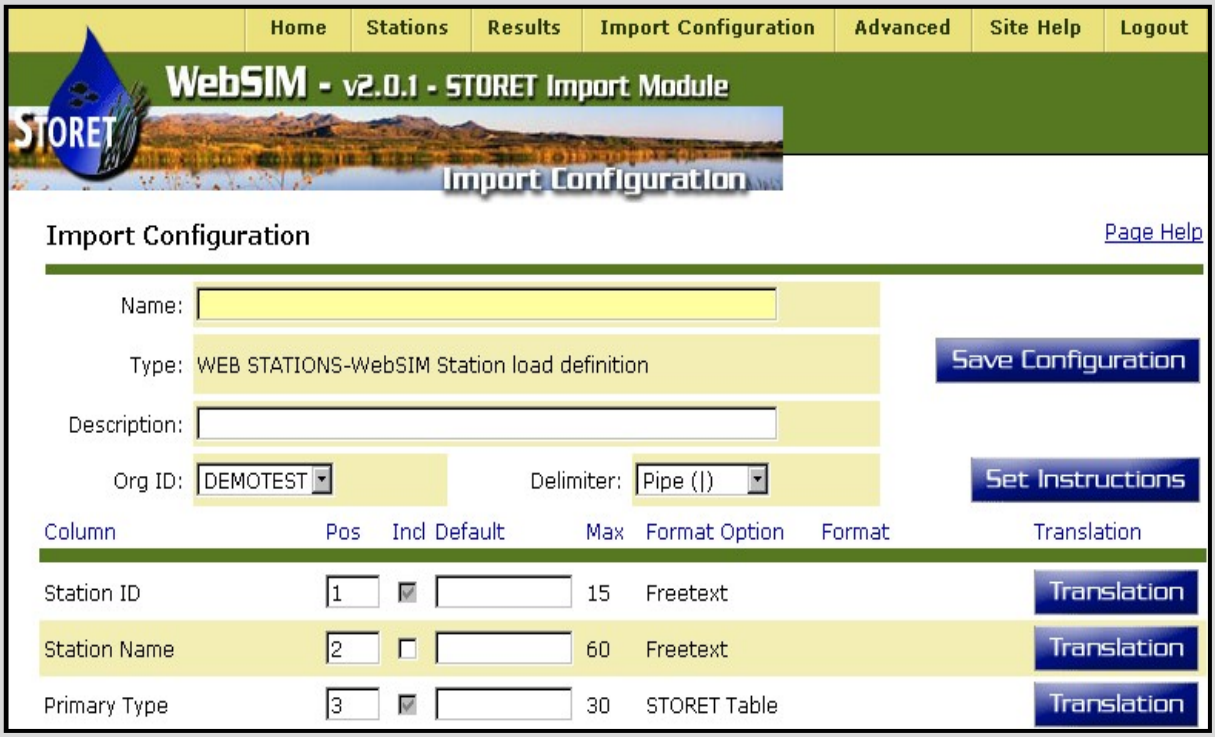


6. This page allows you to specify the type of file that you want to import, either an ASCII flat file or an XML file. (XML files can only be imported for results.)

   Click **Define New Station Import Configuration** under "ASCII Flat File Import Configuration."



You will reach the Import Configuration window.



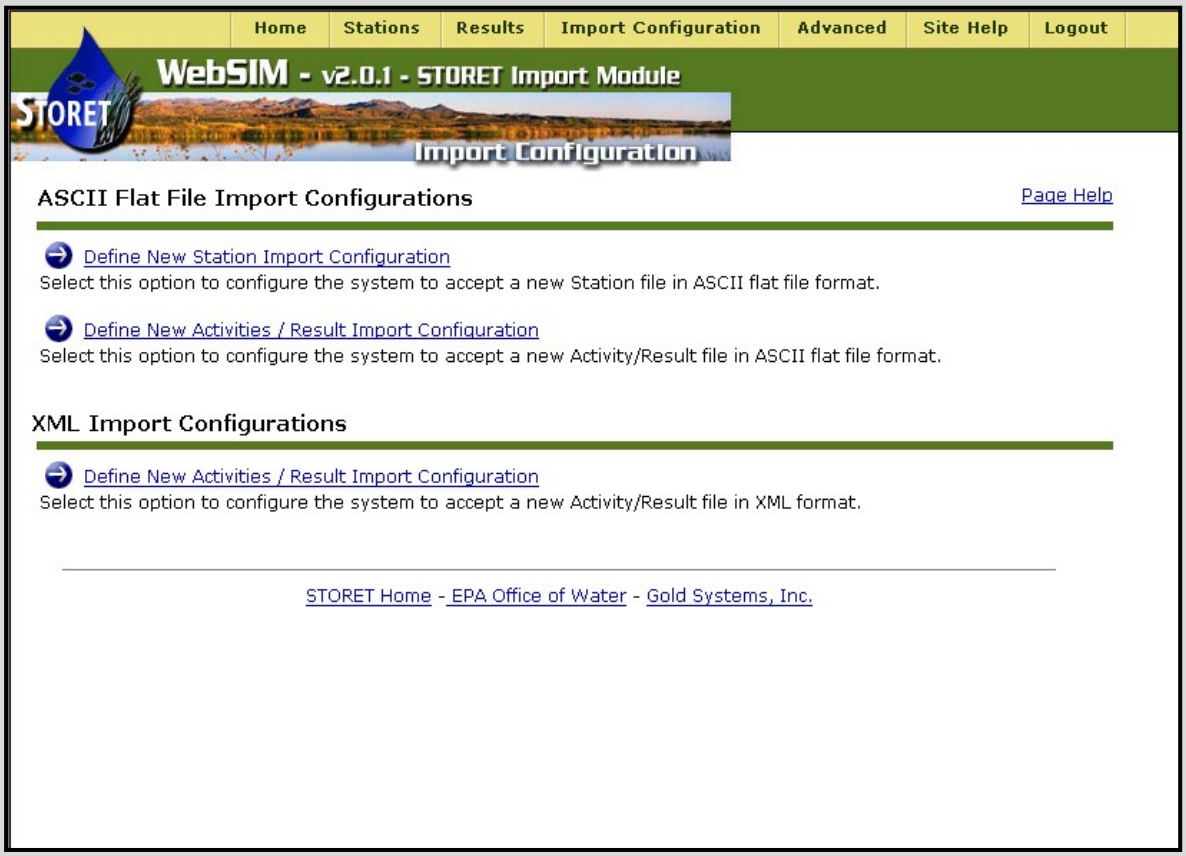
Note: Data elements that are checked and shaded out are required by STORET.

From the "Flat File Import Configuration" menu, you can

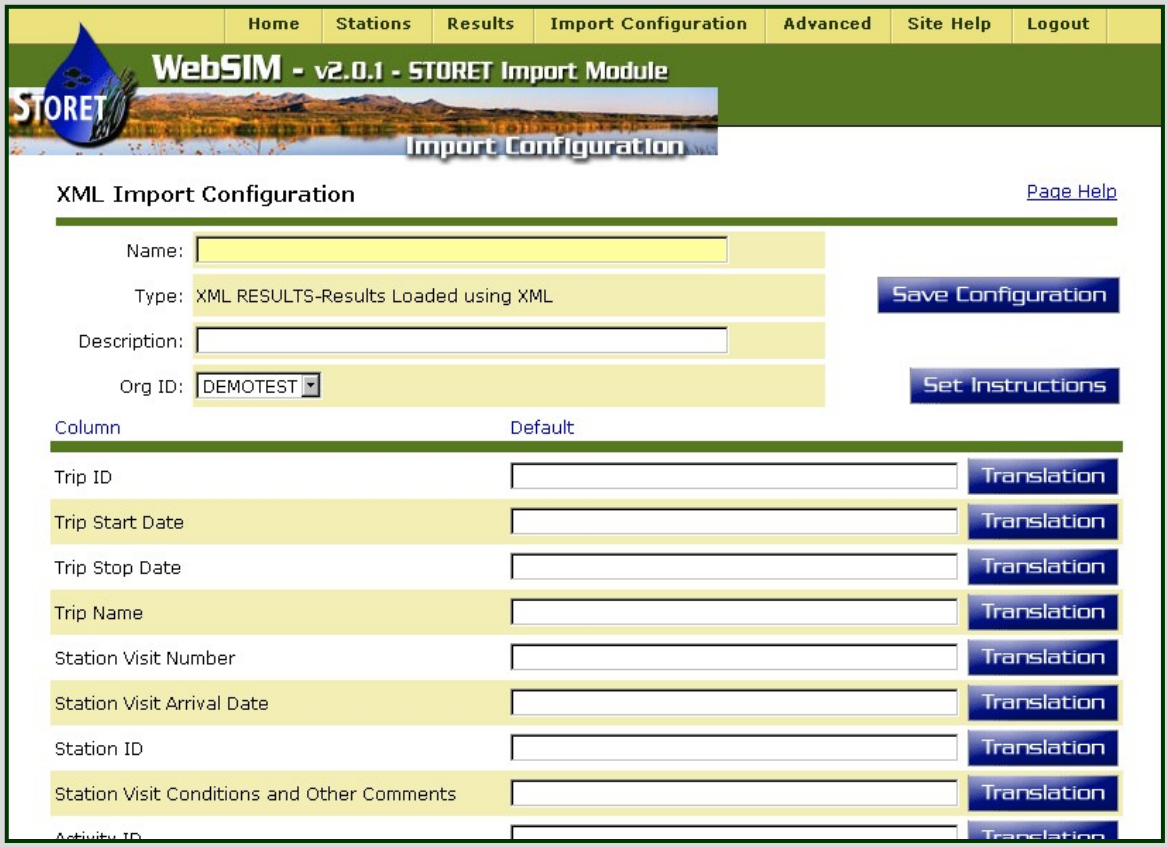
- Provide a name and description for the import configuration
- Define the organization that the data will load into
- Indicate the delimiter that will be used in the text file
- Save the configuration
- Set instructions to review translations and establish general import instructions
- Select the data elements that will be included in the text file by checking the boxes under the "Incl" column header
- Define the position or order in which the data elements will appear in the file
- Establish default values for the data elements in the import file that are left null or blank
- View the format that each row in the import file must adhere to if this configuration is used

- Specify additional formatting for individual data elements (if appropriate)
- Establish translation parameters that change values while a file is being imported.

7. Click the **Back** button in your Web browser.



8. Choose **Define New Activities/Result Import Configuration** under "XML Import Configurations." You can perform essentially the same functions here as listed earlier for flat files. You do not, however, need to indicate the delimiter, because delimiters are not necessary for XML.



9. Click the **Back** button in your Web browser.

10. Provide a name and your ORG ID and click **Save Configuration**. Now you can access the Translation feature.

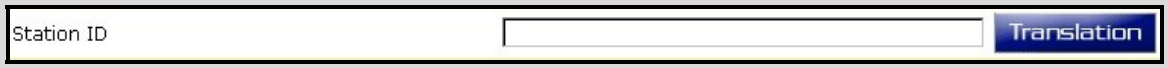
You can use a translation to convert a number or phrase in your current data to one that is accepted by STORET. If you wish to perform a translation after defining an import configuration, click the **Translation** button for the column you want to translate.

The main purpose of the XML import configuration is to set translations. You may want to establish translations to clarify your data. For example, if you have a characteristic in your data, "DO," for dissolved oxygen, you may want to translate "DO" to "Dissolved Oxygen" through the import configuration's Translation feature.

(To set up your XML schema, review the example schema on your MyCDX page. From the MyCDX page, click **eBEACHES: Monitoring**. Next, choose item 2, **Prepare your data**, to see the example.)

11. Click on **Return to Previous** to return to the XML Import Configuration window.

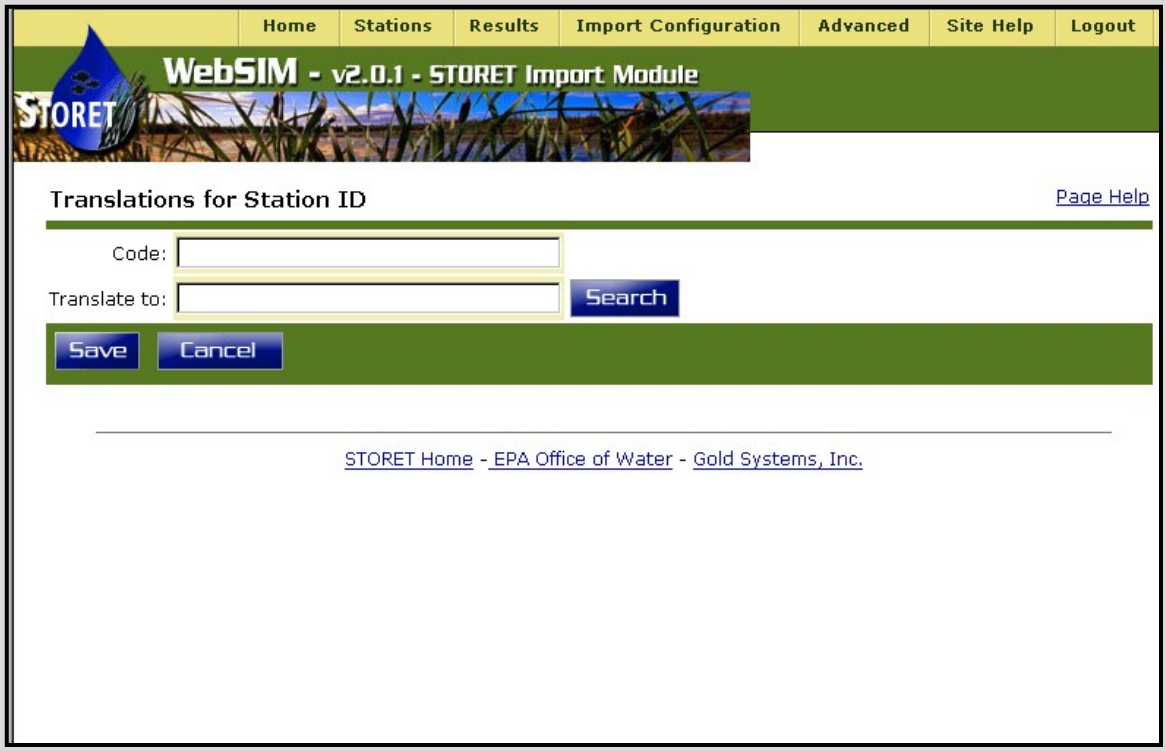
12. Click the **Translation** button next to Station ID.



13. You can select **New Translation** or **Return to Previous.** For this lesson, click **New Translation**.

Translations transform codes in your data set into allowable values in STORET. You can establish a translation by defining the "code" that should be translated and the value that it should be "translated to." A single data element can have several translations.

14. If you are working with test data, just click **Cancel** from this window. If you are using your own data, type in a "code" from your data file and enter the appropriate "translation." Clicking on the **Search** button will display the available allowable values already in WebSIM for this data element. Click **Save** to store this translation with your configuration.



15. Click **Return to Previous** and then click the **Set Instructions** button on the right of the Import Configuration area of the screen menu.

Instructions are typically used to identify data elements that may require translations and to autogenerate ID numbers for data elements such as Trip IDs and Station Visit Numbers. You can use this menu to review translations and establish general import instructions. To change an instruction, click the name of the instruction.

16. Click the **Home** button at the top of the screen to return to the main WebSIM page.

Summary

The purpose of an import configuration is to set up a pattern to translate your data file into STORET using SIM. SIM import configurations can help you set up the order of fields, format data, and perform simple data translations. Ideally, you will set up a few configurations that you will use each time future data are imported to STORET.

In Lesson I, Defining an Import Configuration, you learned how to

- Define a new import configuration
- Save a new import configuration
- Use translations to change values while a file is being imported
- Review general import configuration instructions.

[This page intentionally left blank.]

Lesson II. Using SIM to Migrate Data to STORET

This lesson describes how to

- Import a file into SIM so that it can be validated against the STORET requirements
- Migrate data that pass validation into STORET.

1. Use the toolbar to select **New Import** under "Stations."

2. You must first select an existing import configuration. (Lesson I shows you how to create a new import configuration). Click the drop-down box next to "Configuration" to see a list of import configurations that have been created in SIM. You can review the configuration details by clicking the **Edit Existing Configuration** button. In addition, you can create a new configuration by selecting **Define New Configuration**.

   Select **DEMOTEST--WEBSIM Example - Ocean Station.txt** from the drop-down box beside "Configuration."

3. Then, select an appropriate text file name from the drop-down list beside "Text File Name."

   Select **WEBSIM Example – Ocean Station.txt**.

   Click the **Import File to SIM** button. If you are importing an XML file, be sure to choose an XML import configuration.

SIM checks the file for errors that would prevent the data from being loaded to STORET. From the Import Status menu, you can

- Review the number of records that that SIM imported, the number of those records that had errors, and the number of those records that are ready to be migrated to STORET
- Migrate records to STORET
- Delete an import from SIM. All data remain in SIM until the import is deleted. *Note: You should delete your imports once you are satisfied that your data have been accurately migrated to STORET.*

4. Click **Migrate Records to STORET** and migrate the data you imported to STORET.

*Note: If your import is not successful using this example, it may be because others have already used the Ocean Station example for DEMOTEST and forgot to delete the import. Click the **View Import Errors** link. A table of import errors will appear. Clicking the **Status** button for each row will allow you to delete that import. After no imports remain, you can return to the Import Stations dialog by using the sidebar to navigate to "New Import" under "Stations" and begin the import again.*

Summary

In Lesson II, using SIM to Migrate Data to STORET, you learned how to

- Import files to SIM using a predefined import configuration and text file
- Migrate data into STORET.

[This page intentionally left blank.]

Lesson III. Advanced Functions

This lesson explains how to

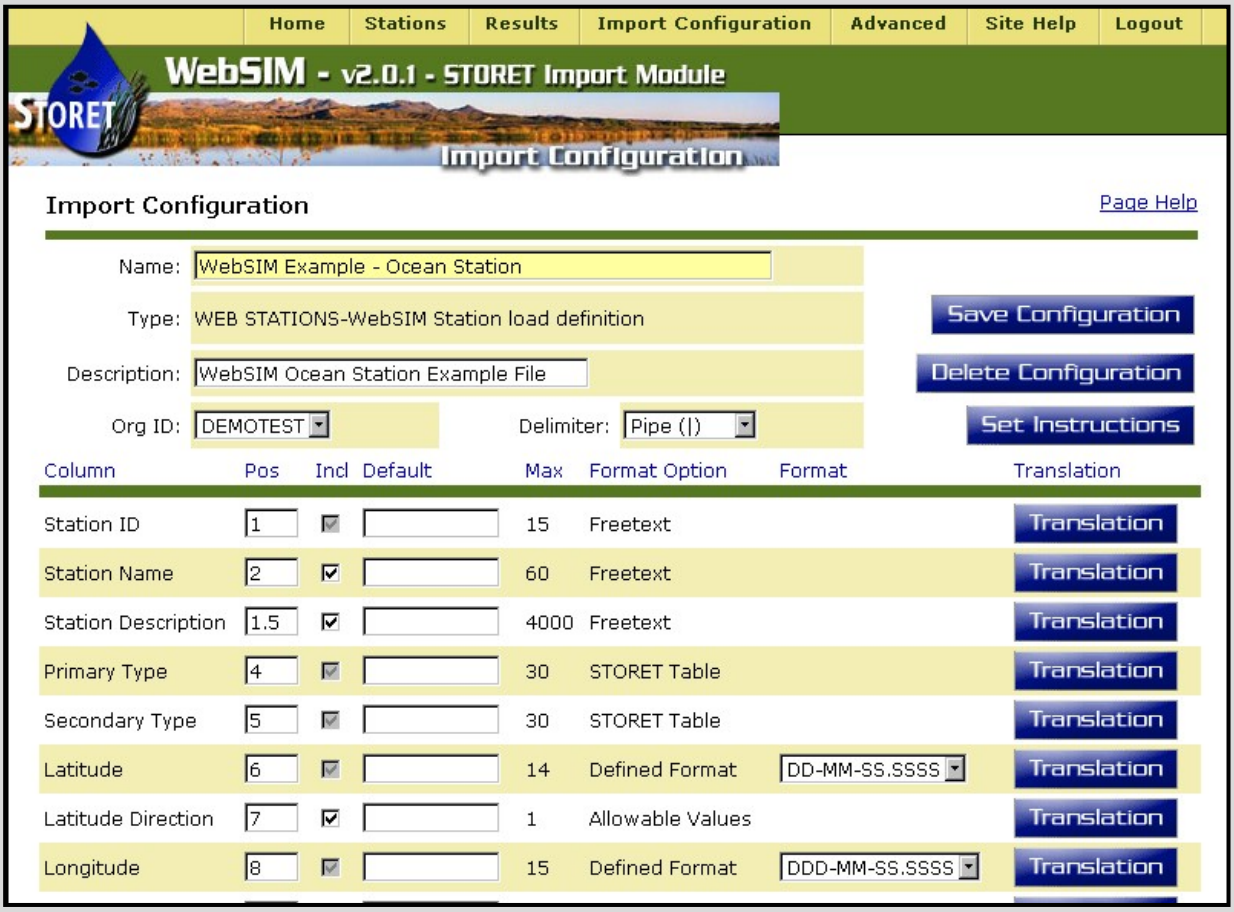
- Reorder field positions in an import configuration
- View the status of import files brought into SIM but not migrated or deleted
- Remove (undo) a migration of data from STORET
- Generate an error report to print or save
- Delete an import from SIM
- View data that have been imported to SIM
- Test and improve system performance.

Reorder field positions in an import configuration

1. From the toolbar, click **Edit Existing** under "Import Configuration." Locate the WebSIM Example – Ocean Station configuration and click the **Edit** button on that row. You may need to click **Next** several times to locate the configuration. You can also sort the configuration list by clicking on any of the column headers – **Configuration Name, Organization ID**, or **Import Type**, which may help you locate the desired configuration more quickly.

You can use the Import Configuration menu to reorder field or column positions.

2. Suppose "Station Description" needs to be in the second position instead of the third. Change the position number from 3 to a number between 1 and 2 (such as 1.5) and click **Save Configuration** to reorder the columns.



View the status of an import in SIM that was not migrated or deleted

In some cases, you may have imported files to SIM but did not migrate the data to STORET. This may have occurred because you had to stop your session or wanted to review a file. This section helps you review previous imports to determine their status. It is important to remember to delete imports after migration.

1. The SIM main screen is always accessible by clicking **Home** on the toolbar. From this page, you can access two import categories: Stations and Results. To review the status of the import that was performed in Lesson II, click **Import Log** under "Stations" on the toolbar menu.

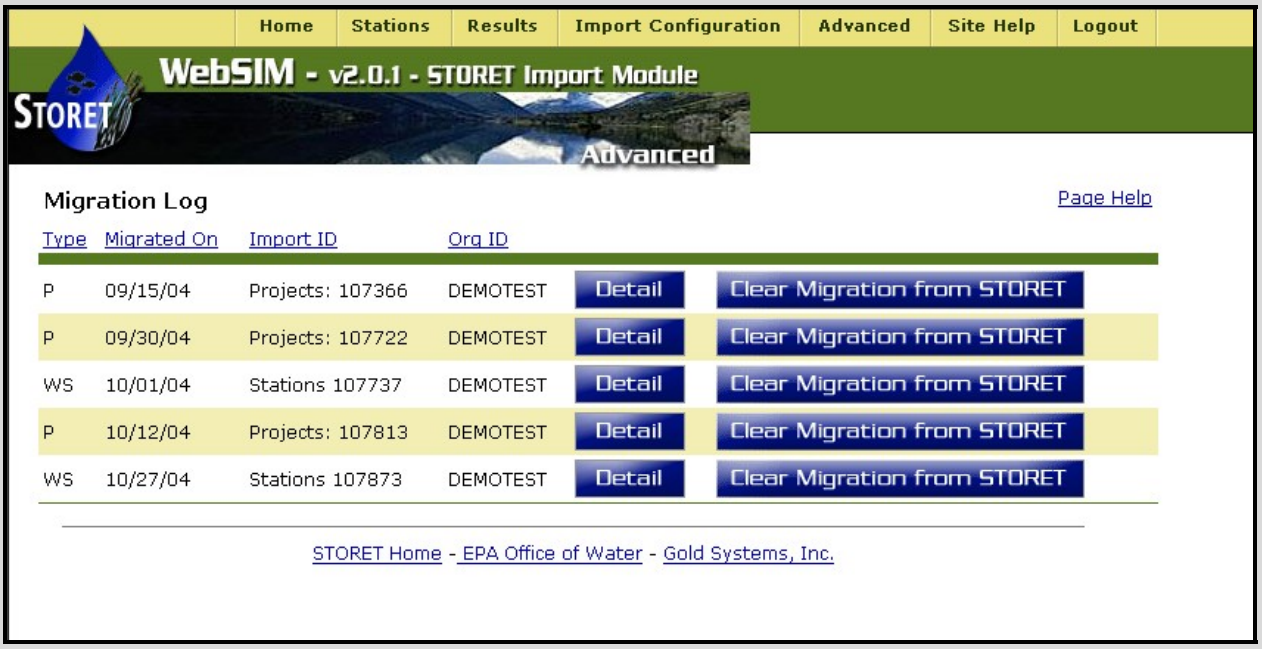
2. Click **Status** at the end of the row that represents the relevant import. You will return to the Import Status window that you saw in step 4 of Lesson II.

Remove a migration of data from STORET

If you migrate a file to STORET and later determine you do not want it in STORET, you can remove that data from STORET. By clearing a migration, you essentially delete the data from STORET.

1. On the toolbar, click **Migration Log** under "Advanced."

2. Locate the record of stations data you just migrated in Lesson II. It should show the date that you migrated the data. Click the **Clear Migration from STORET** button next to this record.

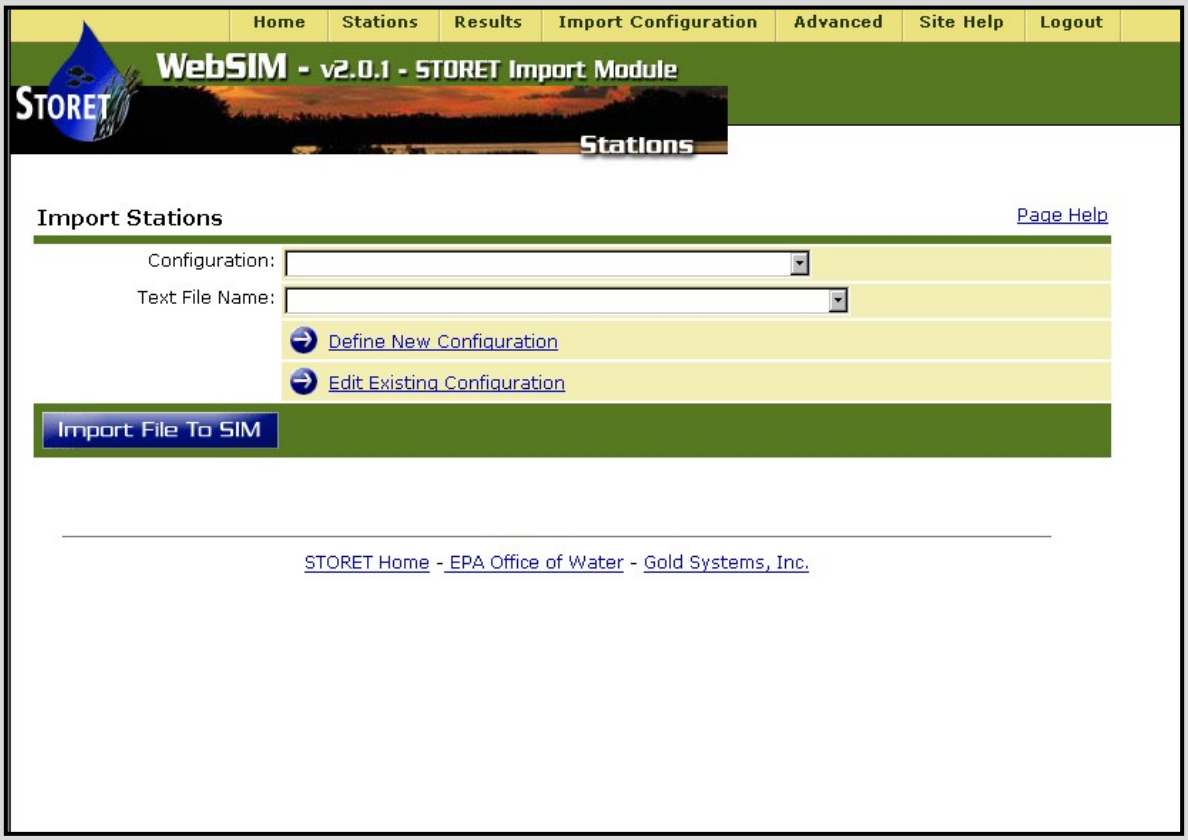


Note: SIM cannot undo a migration once the import has been deleted from SIM.

Generate an error report file to print or save

If your file is imported to SIM with errors, it is important to fix the errors and reimport the file to SIM. Reviewing an error report will help you determine which fields need to be fixed and why.

1. Click **New Import** under "Stations" on the toolbar menu.



2. In order to generate an import with errors, import the same file that you imported in Lesson I.

   Select **DEMOTEST--WEBSIM Example - Ocean Station.txt.**

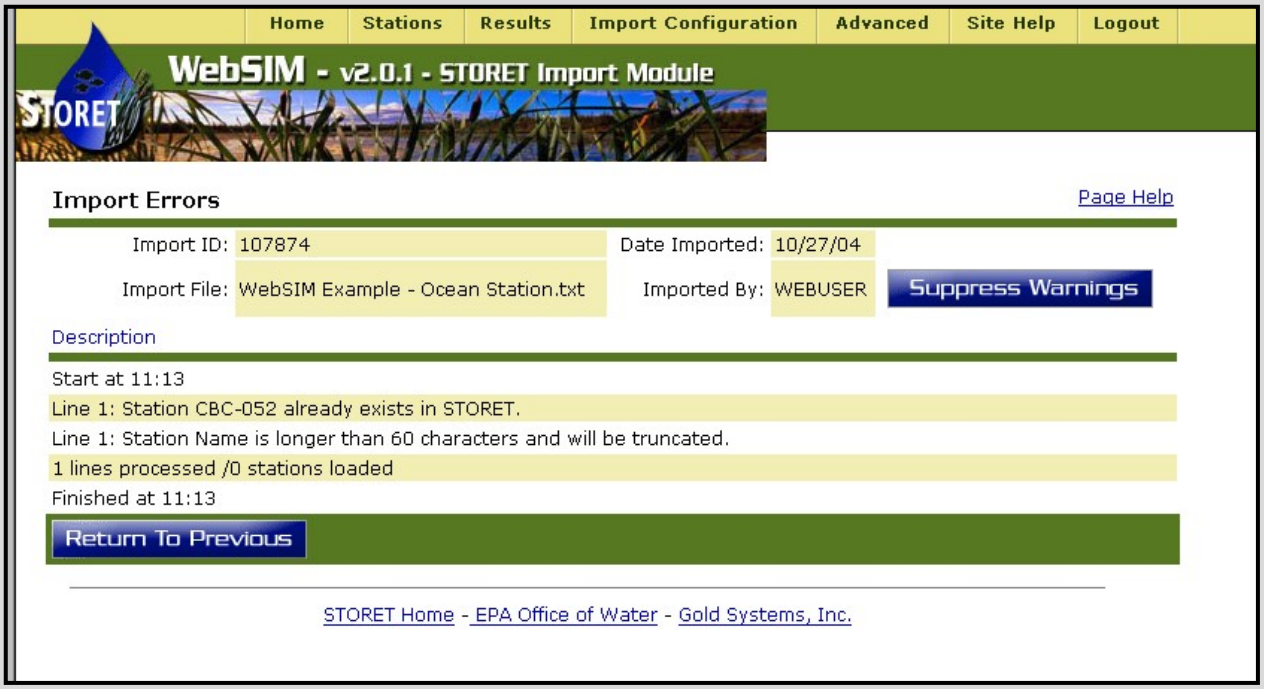
   Select **WEBSIM Example OceanStation.txt**.

   Click **Import File to SIM.**

Note: If there are any errors in your file, you cannot migrate any records to STORET. In addition, if you import a file with errors to SIM, make sure to delete the import file from SIM before you import a new file. If you do not delete the import file with errors, SIM will automatically delete it when you import a new file.

3. Click **View Import Errors**.

4. A report of errors that occurred during import is displayed.



The error message shows you that the station already exists in SIM, since you imported this same file for Lesson 1. In cases where many records are imported, the error messages may be so numerous, that it is preferable to export the error log and read it in a word processing program.

5. Click Return to Previous to return to the Import Status window.

6. The default text file name in WebSIM for exporting errors is unmigrec.txt. You can rename this by clicking in the Export File text box and typing in a different name. For this lesson, just leave the default file name, unmigrec.txt. Click on **Export Records with Errors.**

7. Click on **Return to Import Status**. You will notice that the status window shows that no lines were imported. This is because SIM has exported the line with errors to the unmigrec.txt file.

Delete an Import from SIM

It is important to delete SIM imports once you have finished migrating data into STORET or have fixed import errors and wish to re-import a file already in SIM. If you import a file that is already in SIM, you will receive error messages, as occurred in the previous exercise. If the file in SIM already contains errors and you re-import the file, the new import will overwrite the old one.

1. In the Import Status window, click on Delete Import from SIM to remove the file you just imported in the previous exercise.

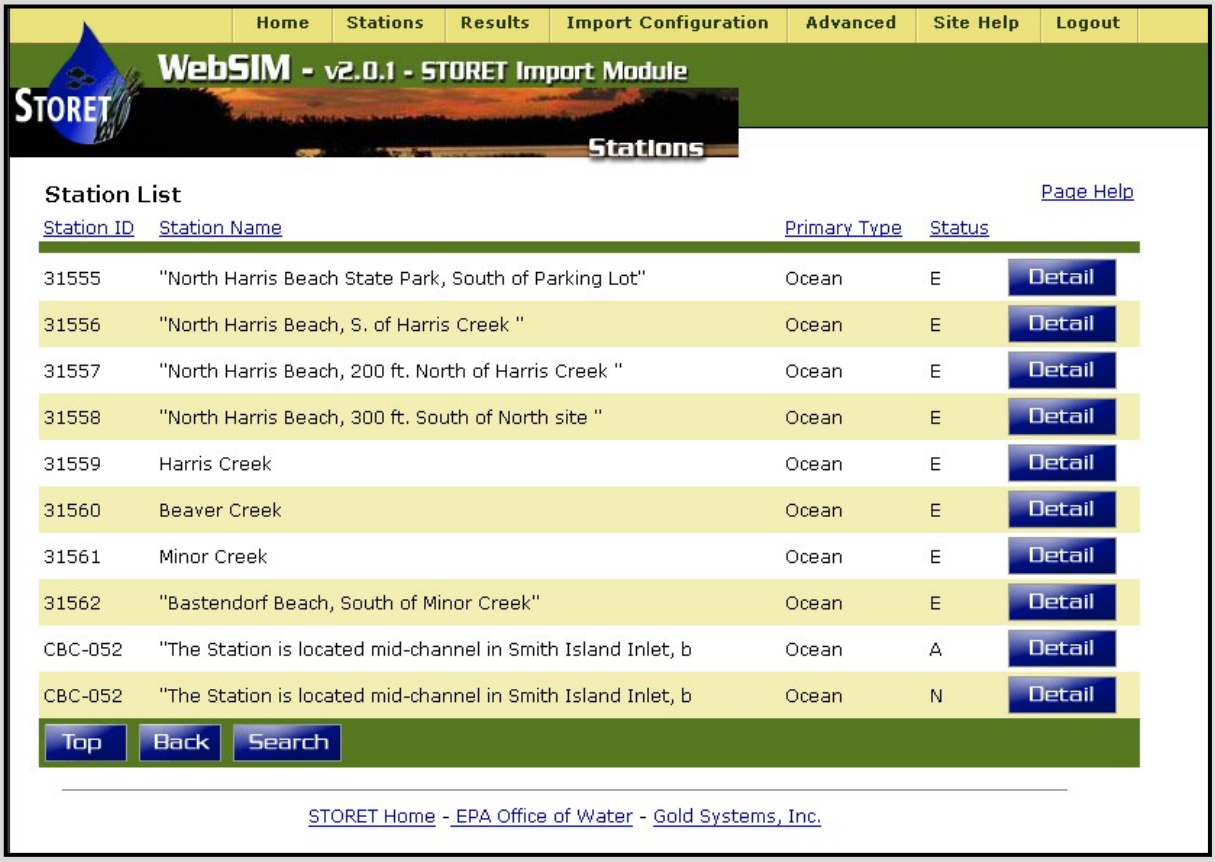
2. Click Continue with Delete when WebSIM asks if you are sure you want to delete this file.

3. After completing the deletion, WebSIM will take you to the Import Log window. You should see a record of the stations you migrated for Lesson 2. It should have the date that you imported the data for the Import Date. Click on **Status**.

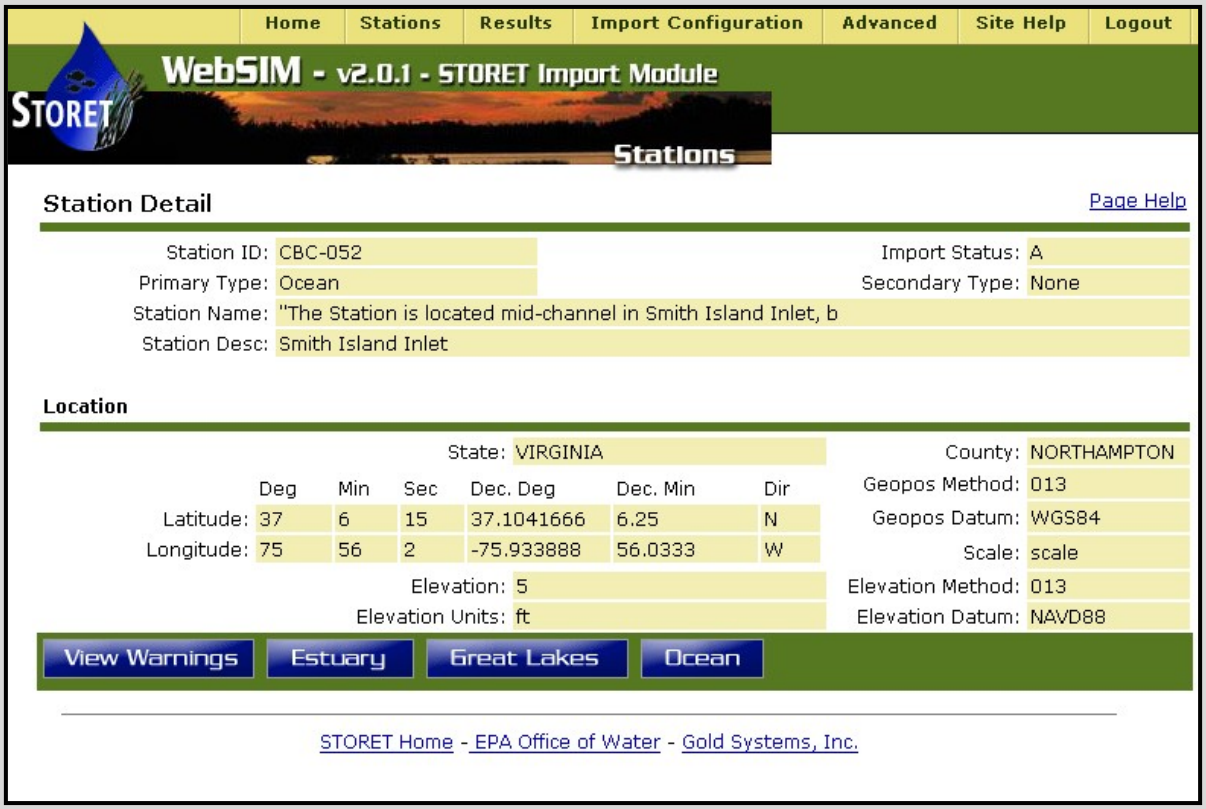
4. In the Import Status window click on Delete Impor from SIM and Continue with Delete in the window that follows.

View data that have been imported to SIM

1. To view data that are in SIM, choose "View" under the Stations and Results menus. Click **View** under "Stations" from the toolbar.



2. Click **Detail** after any individual record to recall and view the data.



Test and improve system performance

- To ensure that the system runs smoothly and will not produce chronic errors, you may want to load a small sample of your data first to test for problems. This may help you avoid errors in multiple records due to an incorrect date format (or other minor error).

- Warnings slow down the import. Warnings are not errors, but general comments from SIM on your import. If you limit the number of warnings that you receive during an import (e.g., latitude direction), then the overall speed of the import will improve. You can do this by clicking on **Suppress Warnings** in the Import Errors window.

Summary

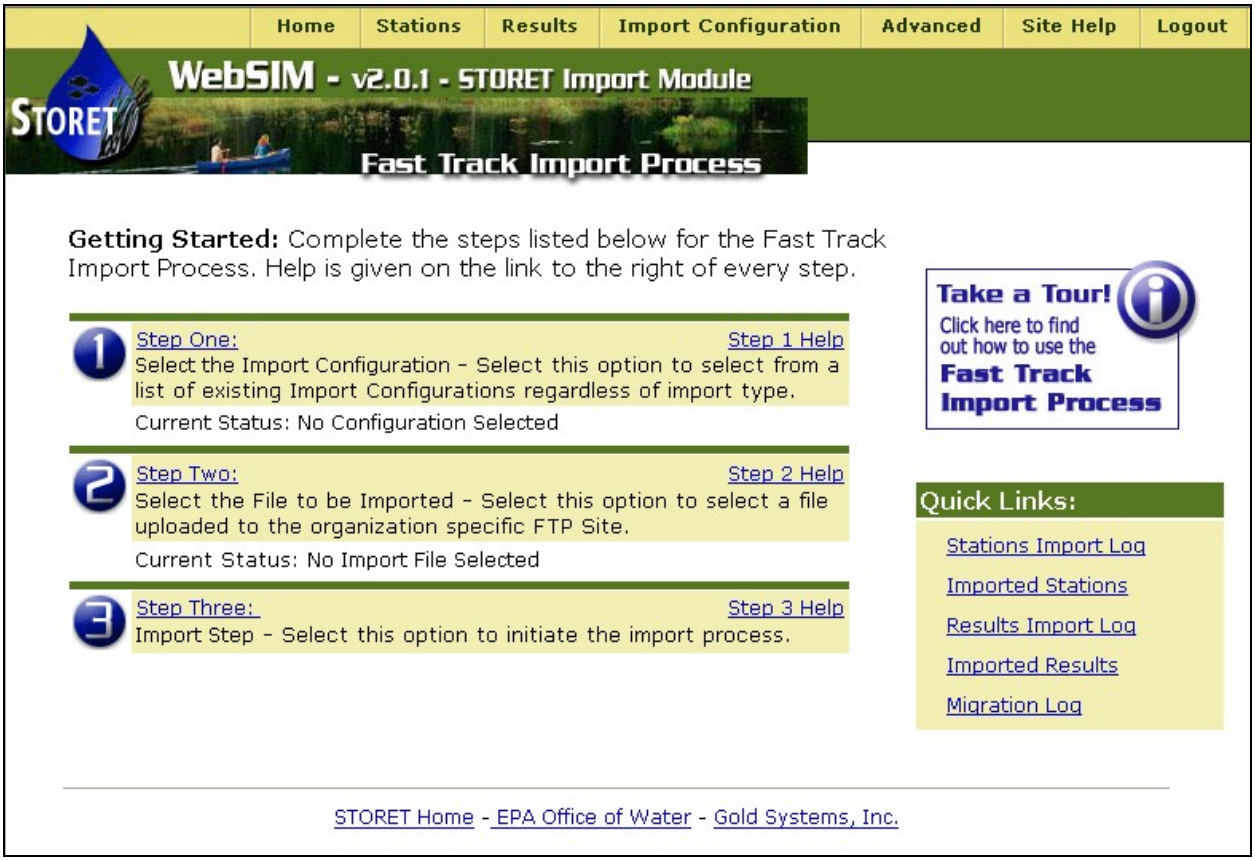
In Lesson III, Advanced Functions, you learned how to

- Reorder field positions in an import configuration
- View the status of import files brought into SIM but not migrated or deleted
- Remove a migration from STORET
- Generate an error report to print or save
- Delete an import fro SIM
- View data that have been imported into SIM
- Test and improve system performance.

[This page intentionally left blank.]

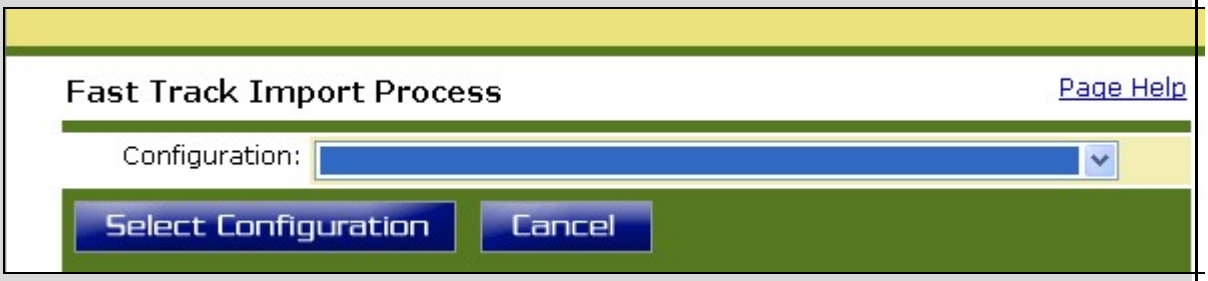
Lesson IV. Using the Fast Track Import Process

The Fast Track Import Process is an easy way to import your data to STORET if there is already an Import Configuration established. The Fast Track Import Process is available from the WebSIM home page.



To use the Fast Track Import Process, follow the steps listed under "Getting Started".

1. When you click **Step One**, you will be able to select an existing Import Configuration from the drop-down list.



2. After you choose the correct configuration, click the **Select Configuration** button.

After you complete each step, the current status will appear in green text below that step.

3. Click **Step Two**. From the drop-down list, choose the file you would like to import. You must have completed the third step of the eBeaches Monitoring Data Road Map (Upload your data to CDX) before your file will appear in this drop-down list.

4. Once you have chosen the correct file to import, click the **Select Import File** button. The green text below Step Two will indicate which file you chose. If this is not the correct file, you may repeat the previous step to select the correct file.

5. Click **Step Three**.

   Here you may choose whether or not to create translations for your import file. Choosing **Create Translations for Station Fields** assists users in identifying errors for the import file. The translations are created for invalid values as the import processes. After the import is complete, you can check the translations page to see what import columns contained invalid values and then complete the translation by including a valid value to translate to.

6. Choose either **Create Translations for Station Fields** or **Don't Create Translations for Station Fields** and then click **Select Instruction**.

7. You will get a message summarizing the data elements you choose during the process. If they are correct, click the **Start Import** button and you will get a message saying "Your import is in progress. Please wait.".

When your import is complete, you will be returned to the WebSIM home page.

This concludes the WebSIM User Guide and Reference Manual.

Appendix A.
Listing of Available STORET Data Elements for Each File Type

This appendix lists data elements that you can enter to describe stations, activities, and results. The tables indicate which elements are required and whether you can enter specific values for each element. These tables may be useful if you are importing flat files to SIM.

For XML import files, review the example schema on your MyCDX page to determine how to set up your XML schema. From the MyCDX page, click **eBEACHES: Monitoring**. Next, choose item 2, **Prepare your data**, to see the example.

Table A-1. Station Descriptions

Data Element	Required	Allowed Values
Station ID	Yes	
Station Name	Yes	
Primary Type	Yes	Yes
Secondary Type	Yes	Yes
Station Description		
Latitude	Yes	
Latitude Direction	Defaults to “N”	Yes
Longitude	Yes	
Longitude Direction	Defaults to “W”	Yes
Geopositioning Datum	Yes	Yes
Geopositioning Method	Yes	Yes
Scale	Conditional	
State	Yes	
County	Yes	
Elevation		
Elevation Units	Conditional	Yes
Elevation Method	Conditional	
Elevation Datum	Conditional	
Ocean Name	Conditional	Yes
Shore Relation	Conditional	Yes
Additional Ocean Name		
Primary Estuary	Conditional	
Secondary Estuary		
Other Estuary		
Great Lake	Conditional	Yes
Additional Great Lake Name		

Table A-2. Activities and Results

Data Element	Required	Allowed Values	Multiple Allowed
Trip ID	Yes	Must Exist in STORET	
Trip Start Date			
Trip Stop Date			
Trip Name			
Station ID	Yes	Must Exist in STORET	
Additional Location Information			
Station Visit Number	Yes		
Station Visit Arrival Date			
Visit Comments			
Project ID	Yes	Must Exist in STORET	Yes. Separate with "\"
Activity ID	Yes		
Medium	Yes	Yes	
Activity Type	Yes	Yes	
Activity Category	Conditional	Yes	
QC Indicator		Yes	
Sample Matrix		Yes	
Replicate Number	Conditional		
Activity Start Date	Yes		
Activity Start Time			
Activity Start Time Zone	Conditional	Yes	
Activity End Date			
Activity End Time			
Activity End Time Zone	Conditional	Yes	
Depth to Activity			
Depth to Activity Units	Conditional	Yes	
Relative Depth		Yes	
Depth Measured From			
Activity Comments			
Sample Collection Procedure ID	Conditional	Must Exist in STORET	
Gear ID	Conditional	Yes	
Gear Configuration ID		Must Exist in STORET	
Gear Deployment Comments			
Sample Preservation, Transport & Storage ID		Must Exist in STORET	
Sample Transport and Storage Comments			
Detection Condition		Yes	
Characteristic Group ID	Conditional	Must Exist in STORET	
Characteristic Row ID	Conditional	Must Exist in STORET	
Characteristic Name	Conditional	Yes	
Result Value	Conditional	Yes*	
Result Value Units	Conditional	Yes	
Result Status	Defaults to "F"	Yes	
Sample Fraction	Conditional	Yes	
Statistic Type		Yes	
Value Type	Defaults to "Actual"	Yes	
Duration Basis		Yes	
Temperature Basis		Yes	
Weight basis		Yes	
Result Comment			
Laboratory ID		Must Exist in STORET	
Field/Lab Procedure	Conditional	Must Exist in STORET	

Data Element	Required	Allowed Values	Multiple Allowed
Field/Lab Procedure Source	Conditional	Yes	
Laboratory Certified		Yes	
Laboratory Batch ID			
Analysis Date			
Analysis Time			
Analysis Time Zone	Conditional	Yes	
Detection Limit			
Detection Limit Unit	Conditional	Yes	
Detection Limit Comment			